CS 322 Prelim 1

Tuesday 20 February 2006—90 minutes

Problem 1: Linear Systems (30 pts)

Here is a linear system at some intermediate step of one of the linear system solution methods we've looked at.

$$\begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 10^{-10} & 1 \\ 0 & 0 & -2 & 1 \end{bmatrix}$$

Assume that there is nothing special about the original matrix (no special pattern of zeros, for example).

1. What method are we looking at? How did you tell?

   Answer: We are looking at Gauss-Jordan elimination. I can tell because the matrix is being transformed to the identity matrix, and in L-U factorization the matrix is only transformed to be upper triangular.

2. What iteration is the algorithm about to start?

   Answer: It is about to begin the third iteration

3. Which element should be used as the pivot in this iteration? Why?

   Answer: -2, in the fourth row, because $|-2| > |10^{-10}|$ and this will allow us to avoid dividing by an extremely small number.

4. Compute the matrix after the iteration is complete.

   Answer:

$$\begin{bmatrix} 1 & 0 & 0 & 5.5 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & -0.5 \\ 0 & 0 & 0 & 1 + 5\mathrm{e}^{-11} \end{bmatrix}$$

5. What data other than this matrix would the algorithm need to modify during the iteration, and how?

   Answer: The algorithm would need the right hand side vector or matrix that we are solving for, i.e. $\mathbf{b}$ from $\mathbf{Ax} = \mathbf{b}$. It would need to perform all operations on the right hand side that it has done on the matrix **A** – it should swap rows 3 and 4, and perform all the row scales and row additions that were done to form the matrix in part (4).

Problem 2: Problem classification (28 pts)

Consider the following numerical problems:

1. Solving $\mathbf{MQ} = \mathbf{P}$ for $\mathbf{M}$, where $\mathbf{Q}$ and $\mathbf{P}$ are both 4×4.

   Answer: (a). After taking the transpose on both sides, we have a straightforward linear system $\mathbf{Q}^T\mathbf{M}^T = \mathbf{P}^T$ that we can solve for $\mathbf{M}^T$, taking the transpose at the end to get $\mathbf{M}$.

2. Finding an equation of the form $y = ax^3 + bx^2 + cx + d$ to fit 4 points (x_i, y_i).

   Answer: (a). A cubic equation is uniquely defined by its values at 4 distinct points. We can set up a linear system to find the coefficients (a, b, c, d).

3. Finding the value of some function $f(x)$ at x', given values $y_1 = f(x_1), y_2 = f(x_2), ...$ where $x_1 < x_2... < x_n$ and $x_1 \leq x' \leq x_n$.

   Answer: (c). We can use the values at the two neighboring points in order to linearly interpolate, or we can use the points over a wider region to get a higher order fit.

4. Finding values for n variables x_i that best satisfy n linear equations.

   Answer: (a). Definition of a system of linear equations

5. Finding an equation of the form $y = ae^{bx} + c$ to fit five points (x_i, y_i).

   Answer: (d). Because of the b and c terms, we cannot convert this into an equation where each term is linear in the unknowns (a, b, c).

6. Finding a linear transformation of the plane (a 2×2 matrix) to match 3 pairs of points.

   Answer: (b). There is no translation in a linear transformation of the 2D plane, so this is an overconstrained system. One other way of seeing this is to note that the linear transformation is 2×2, while our left hand and right hand side will be both 3×2 – we have more equations than unknowns.

7. Finding a rotation of the plane to match 3 pairs of points.

   Answer: (d). Although applying a rotation is a linear transformation, finding the rotation of the 2D plane to fit 3 points results in nonlinear equations. In particular, the 2×2 rotation matrix has nonlinear dependencies between the entries, and none of the methods discussed will work well on this problem.

8. Finding values for n variables x_i that best satisfy m linear equations, $m > n$.

   Answer: (b). Definition of an overconstrained system.

Match each problem to the one of the following methods that is **best suited** to solving the problem, and give a one-sentence explanation for each.

(a) solving a system of linear equations;

(b) solving a linear least squares system;

(c) interpolation;

(d) none apply.

Problem 3: Least squares (21 pts)

Here are two least squares problems in the standard form $\mathbf{A}\mathbf{x} \approx \mathbf{b}$:

$$1: \quad \mathbf{A} = \begin{bmatrix} 0 & -1 \\ 2 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix} \quad ; \quad \mathbf{b} = \begin{bmatrix} -3 \\ 7 \\ -3 \\ -1 \end{bmatrix}$$

$$2: \quad \mathbf{A} = \begin{bmatrix} 1 & 5 \\ 2 & 4 \\ 3 & 3 \\ 4 & 2 \\ 5 & 1 \end{bmatrix} \quad ; \quad \mathbf{b} = \begin{bmatrix} -3 \\ 3 \\ -3 \\ 9 \\ 9 \end{bmatrix}$$

1. What does it mean for a vector $\mathbf{x}$ to be the solution to a linear least squares system in this form?

   Answer: One meaning is that it satisfies $\min \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$. Another is that it satisfies the system of normal equations $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$.

2. Is $\mathbf{x} = \begin{bmatrix} 2 & 3 \end{bmatrix}^T$ the solution to system 1? Explain how you can tell.

   Answer: Yes. When we plug it into the system of normal equations $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$, we see that $\mathbf{x}$ is a solution to this linear system, and so it must be the solution to the linear least squares problem.

3. Is $\mathbf{x} = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$ the solution to system 2? Explain how you can tell.

   Answer: No. When we plug it into the system of normal equations, we see that $\mathbf{x}$ is not the solution to the linear system, and so it cannot be the solution to the linear least squares problem. Additionally, $\begin{bmatrix} 2 & -1 \end{bmatrix}^T$ has a smaller residual $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$, so $\begin{bmatrix} 2 & 1 \end{bmatrix}^T$ is not the minimal residual.

Problem 4: Snowfall (21 pts)

After the recent snowfall in Ithaca, we gathered a bunch of data from friends around the area who measured the depth of the snow in their back yards. That is, for a set of n locations (x_i, y_i) (denoting longitude x and latitude y), we have a measurement of the snowfall s_i received at that point. Now we wish to find a model to predict snowfall based on position.

1. Suppose we think that snowfall all across Ithaca for the most recent storm can be predicted by a single linear function in x and y. Set up the system that you would need to solve to find this linear function. Be specific about what variables you are solving for and the structure of your matrices.

   Answer: We are trying to find an affine linear function $ax + by + c = s$ that best fits the provided data. We can solve this by solving the corresponding linear fitting problem:

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & & \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

   for a, b, c, which are the coefficients of the linear function we are looking for.

2. Suppose we think a single quadratic function in x and y would be a better fit. Can we still compute it with linear fitting? Why or why not?

   Answer: Yes, we can. We are now looking for a function $ax^2 + by^2 + cxy + dx + ey + f = s$ that best matches the data. Although it is quadratic in x and y, it is still linear in the unknowns, namely the coefficients. We can set up our linear fitting problem in a similar fashion as before, except we will have six unknowns, and our left hand side matrix will be $n \times 6$ in size.

3. Suppose our data was structured in such a way that we could easily use some form of linear interpolation between the closest sample points, and that we had a large number of well-spaced sample points.

   (a) Give a suitable constraint on the structure of the input data that will make this operation easy.

   Answer: If our input data were arranged in an aligned grid, then we can easily linearly interpolate between neighboring points. Since we gathered this data from our friends it is unlikely that they are grid aligned, but we can perform scattered data interpolation in order to compute the values on grid aligned points.

   (b) Interpolation might produce predictions that are more or less accurate than the functions you fit in (1) and (2). Give two arguments, based on different assumptions, the first claiming that interpolation is better, and the second claiming that a low-order fit is better. Be specific about your assumptions.

Answer: Interpolation would probably be better if there were significant variation due to local geographic features, for instance the gorges receiving much more snowfall than downtown. This local variation would not be captured by a single linear model that tries to approximate all of the data.

The linear model would probably be better if the snowfall was only dependent on large scale geographic features (distance from the lake) and we lacked data points in certain regions. In this case, we might feel more confident about our predictions in the area which lacks data than if we tried to use local interpolation to project predictions into these regions.